



# Creating a public resource for protein allergenicity assessment: The COMPARE Database

November 9, 2017



- A comprehensive and up to date collection of peer reviewed allergen protein sequences. Builds on a > 10 year database of peer review science.
- Includes known allergens from 2016 and previous years (listed in existing public database AllergenOnline V16) and *new* allergens identified through the COMPARE process.
- Available since February 03, 2017 at [www.comparedatabase.org](http://www.comparedatabase.org) (website hosted and maintained by HESI)
- Updated annually




[About](#) [Database](#) [Timeline](#) [FAQ](#) [Contact Us](#)

## Compare Database

[[Search Database](#) | [COMPARE 2017 FASTA sequences \(PDF\)](#) | [COMPARE 2017 FASTA sequences \(Text\)](#)]

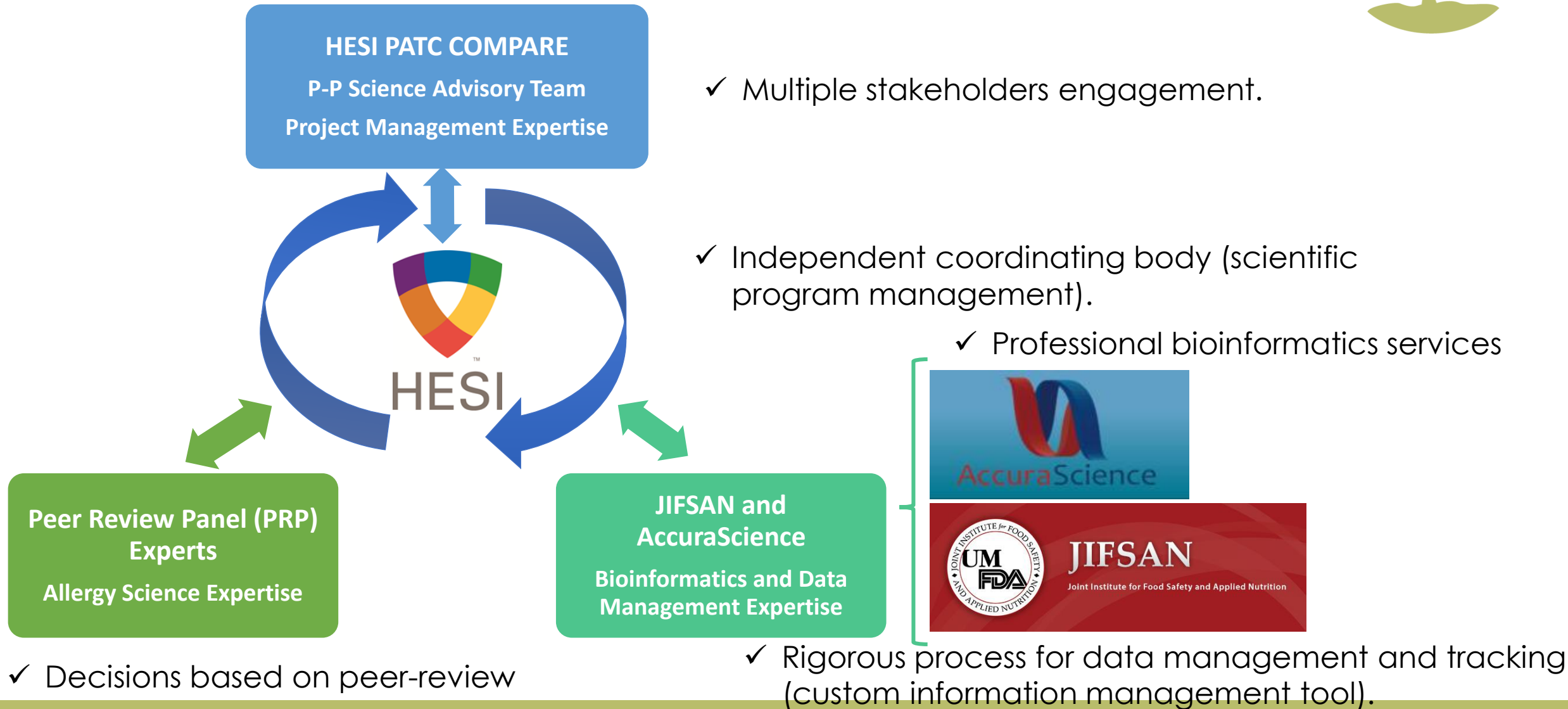
All sequence records originated from the National Center for Biotechnology Information (NCBI) database. As of 02 February 2017, all sequence records are valid.

[Download Data](#)

SPECIES ▲	COMMON NAME	GI #	ACCESSION #	LENGTH	YEAR ADOPTED	
Acarus siro	Mite	118638268	ABL09307.1	131	2009	<a href="#">View Details</a>
Acarus siro	Mite	118638278	ABL09312.1	517	2009	<a href="#">View Details</a>
Actinidia arguta	Hardy Kiwi	441482362	AGC39172.1	213	2014	<a href="#">View Details</a>
Actinidia arguta	Hardy Kiwi	441482364	AGC39173.1	213	2014	<a href="#">View Details</a>
Actinidia arguta	Hardy Kiwi	441482366	AGC39174.1	213	2014	<a href="#">View Details</a>
Actinidia chinensis	Kiwi	68064399	P83958.1	20	2007	<a href="#">View Details</a>
Actinidia chinensis	Kiwi	190358935	P00785.4	380	2009	<a href="#">View Details</a>
Actinidia chinensis	Kiwi	281552896	CAM31908.1	159	2011	<a href="#">View Details</a>
Actinidia chinensis	Kiwi	378548410	P85204.1	15	2013	<a href="#">View Details</a>
Actinidia chinensis	Kiwi	441482354	AGC39168.1	213	2014	<a href="#">View Details</a>
Actinidia chinensis	Kiwi	441482370	AGC39176.1	225	2014	<a href="#">View Details</a>
Actinidia deliciosa	Kiwi	100		109	2016	<a href="#">View Details</a>
Actinidia deliciosa	Kiwi	15984	CAA34486.1	380	2007	<a href="#">View Details</a>
Actinidia deliciosa	Kiwi	166317	AAA32629.1	380	2007	<a href="#">View Details</a>
Actinidia deliciosa	Kiwi	40807635	AAR92223.1	116	2007	<a href="#">View Details</a>

# Rationale for building COMPARE @ HESI:

## 1) Key components of the HESI public-private partnership.



# Rationale for building COMPARE @ HESI:

2) Good Coordination allows multiple sourcing of allergen candidates



GenBank



*Bioinformatics automated  
filtering process*

IUIS



Literature Dbases



*Targeted search strategy*

Allergen  
Candidates  
List

Peer  
Review

New  
allergens

**Goal:** enhance confidence (no allergens left behind)

Add to existing  
database:  
**COMPARE  
Update**

# Rationale for building COMPARE @ HESI:

## 3) Enhancing transparency with detailed documentation




### What has been done:

- COMPARE overview brochure: at initiation of the project (Q1 2016).
- Detailed process development described in the COMPARE website: within 6 months of the database public release (COMPARE' 2017).  
<http://comparedatabase.org/process-development/>
- Announcement letter sent to stakeholders and published on JIFSAN's website.

### In preparation:

- Peer-review publication





Comprehensive Protein Allergen Resource

**COMPREHENSIVE PROTEIN ALLERGEN RESOURCE**

Program Overview – February 2016


#### ABOUT HESI

HESI, the Health and Environmental Sciences Institute, is a non-profit, global scientific organization dedicated to developing science for a safer, more sustainable world.

HESI's mission is to collaboratively identify and help to resolve global health and environmental challenges through engagement of scientists from academia, government, industry, NGOs, and other strategic partners.

#### OVERVIEW

The COMPARE Database, to be made publicly available in January 2017, will be a transparent resource for identification of protein sequences that are known or putative allergens. The COMPARE database will meet needs for allergy safety assessment via an annual updating process that 1) captures new listings of allergen sequences, 2) filters out non-allergen sequences, 3) identifies published literature linked to the identified potential allergen sequences, and 4) verifies that the newly identified sequences have clinical, published evidence of allergenicity based on standardized criteria.



#### PROGRAM STRUCTURE

The COMPARE database is a collaborative HESI program that combines programmatic support from the Joint Institute for Food Safety and Nutrition – [www.jifsan.org](http://www.jifsan.org) at the University of Maryland - with financial support from private sector partners. Additional in-kind support from multiple academic and government partners is engaged throughout the process. A public-private HESI collaborative team will publicly document all database design and search algorithm decisions. An exclusively public-sector expert panel will review the output from a high-throughput bioinformatic pipeline against peer-reviewed publications. The expert panel will then select sequences for inclusion in the database and document their rationale for inclusion.

#### TIMELINE

The COMPARE database will be released to the public by January, 2017. To enable this timing, a steering team will convene in early 2016 to ensure timelines, process, and resources are aligned. Sequence searching will be conducted in second quarter 2016 and peer-review in third quarter 2016 before final assembly and release.

# Conclusion: overall program design and components enable long term, high quality and stable support.



Advisory Team	Bioinformatics and Data Communication	Program Coordination	Expert Peer Review Panel	Participating Organizations
<ul style="list-style-type: none"> <li>• 20 Public/Private stakeholders.</li> <li>• General program oversight, QC, user feedback.</li> <li>• Expert knowledge</li> <li>• No influence on decisions regarding sequence inclusion or exclusion.</li> </ul>	<ul style="list-style-type: none"> <li>• <b><u>AccuraSciences:</u></b> keyword filtering and algorithm design; sources information for database.</li> <li>• <b><u>University of Maryland / JIFSAN:</u></b> <ul style="list-style-type: none"> <li>• Literature search and procurement.</li> <li>• Data communication.</li> <li>• Review Tool.</li> </ul> </li> </ul> 	<ul style="list-style-type: none"> <li>• Overall program management.</li> <li>• Coordination between partners, PRP assembly and management.</li> <li>• Database 'home' and maintenance.</li> </ul> 	<ul style="list-style-type: none"> <li>• Invited public sector allergy experts.</li> <li>• Define inclusion /exclusion criteria.</li> <li>• Review of sequences and references / publications.</li> <li>• Provide final decisions on new sequences to include in database</li> </ul>	<ul style="list-style-type: none"> <li>• <b>Public Sector:</b> <ul style="list-style-type: none"> <li>• Universities of Amsterdam</li> <li>• University of Maryland / JIFSAN</li> <li>• U.S. EPA</li> <li>• U.S. FDA</li> </ul> </li> <li>• <b>Private Sector:</b> <ul style="list-style-type: none"> <li>• BASF</li> <li>• Bayer CropScience</li> <li>• Dow</li> <li>• DuPont Pioneer</li> <li>• KWS</li> <li>• Monsanto</li> <li>• Syngenta</li> <li>• Vilmorin &amp; Cie</li> </ul> </li> </ul>





# Who will use the COMPARE Database?



## Who?

- **Target audiences:** regulatory agencies as well as the public (lay person(s) and professionals)

## Rationale?

- **Truly public resource:** any/all stakeholders or users of the database have equal and full access. Ensures consistency and clarity.

## How?

- **Sequence comparisons:** these can be made between biotech proteins and known allergens – COMPARE allows these bioinformatic comparisons.
- **Basic research** on types of allergens, sequence information and literature associated with clinical relevance for including specific allergens.
- **Staying updated** with new allergens identified each year (all users).

**Access to a transparent and consensus-based allergen database is a key aspect for supporting public safety.**



## Visit:

- ✓ HESI PATC website: committee fact sheet, events and publications:  
<http://hesiglobal.org/protein-allergenicity-technical-committee-patc/>
- ✓ COMPARE Database: [www.comparedatabase.org](http://www.comparedatabase.org)



[www.hesiglobal.org](http://www.hesiglobal.org)



